SEQUENCE LISTING

<110> Romeo, Tony
Wang, Xin

<120> METHODS FOR POLYSACCHARIDE ADHESION SYNTHESIS MODULATION

<130> 14233.10USU1

<140> US 10/675,226
<141> 2003-09-29

<150> US 60/414,352
<151> 2002-09-30

<160> 9

<170> PatentIn version 3.1

<210> 1
<211> 2460
<212> DNA
<213> Escherichia coli

<400> 1
atgtattcaa gtagcagaaa aaggtgcccg aaaaccaaat gggctttgaa acttcttact 60

gccgcatttt tagcagcgag tcccgcggcg aagagtgctg ttaataacgc ctatgatgca 120

ttgattattg aagctcgcaa gggtaatact cagccagctt tgtcatggtt tgcactaaaa 180

tcagcactca gcaataacca aattgctgac tggttacaga ttgccttatg ggccgggcaa 240

gataaacagg ttattaccgt ttacaaccgc taccgtcatc agcaattacc agcgcgtggt 300

tatgcagctg tcgccgtcgc ttatcgtaac ctgcaacaat ggcaaaactc gcttacactg 360

tggcaaaagg cgctctctct ggagccgcaa aataaggatt atcaacgggg acaaatttta 420

accctggcag atgctggtca ctatgatact gcgctggtta aacttaagca gcttaactct 480

ggagcaccgg acaaagccaa tttactcgca gaagcctata tctataaact ggcggggcgt 540

catcaggatg aattacgggc gatgacagag tcattacctg aaaatgcatc tacgcaacaa 600

tatcccacag aatacgtgca ggcattacgt aataatcaac ttgctgccgc gattgacgat 660

gccaatttaa cgccagatat tcgcgctgat attcatgccg aactggtcag actgtcgttt 720

atgcctacgc gcagtgaaag tgaacgttat gccattgccg atcgcgccct cgcccaatac 780

gctgcattag aaattctgtg gcacgataac ccagaccgca ctgcccagta ccagcgtatt 840

caggttgatc atcttggcgc gttattaact cgcgatcgtt ataaagacgt tatttctcac 900

tatcagcgat taaaaaagac ggggcaaatt attccgccct gggggcaata ttgggttgca 960

tcggcttatc tcaaagatca tcagccgaaa aaagcacagt caataatgac cgagctcttt 1020

tatcacaagg agaccattgc cccggattta tccgatgaag aacttgcgga tctcttttac 1080

agccacctgg agagtgaaaa ttatccgggc gcgctaactg tcacccaaca taccattaat 1140

acttcgccgc ctttccttcg gttaatgggc acgcctacga gcatcccgaa tgatacctgg 1200

ttacaggggc attcgtttct ctcaaccgta gcaaaatata gtaatgatct tcctcaggct 1260

gaaatgacag ccagagagct tgcttataac gcaccaggaa atcagggact gcgcattgat 1320

tacgcgagtg tgttacaagc ccgcggttgg cctcgtgcag cagaaaatga attaaaaaaa 1380

gcagaagtga tcgagccacg taatattaat ctggaggttg aacaagcctg gacagcatta 1440

acgttacaag aatggcagca ggcagctgtc ttaacgcacg atgttgtcga acgtgaaccg 1500

caagatcccg gcgttgtacg attaaaacgt gcggttgatg tacataatct tgcagagctt 1560

cgtatcgctg gctcaacagg aattgatgcc gaaggcccgg atagtggtaa acatgatgtc 1620

gacttaacca ccatcgttta ttcaccaccg ctgaaggata actggcgcgg ttttgctgga 1680

ttcggttatg ccgatggaca atttagcgaa ggaaaaggga ttgttcgcga ctggcttgcg 1740

ggtgttgagt ggcggtcacg taatatctgg ctcgaggcag agtacgctga acgcgttttc 1800

aatcatgagc ataaacccgg cgcgcgcctg tctggctggt atgattttaa tgataactgg 1860

cgtattggtt cgcaactgga acgcctctct caccgcgttc cattacgggc aatgaaaaat 1920

ggtgttacag gcaacagtgc tcaggcttat gttcgctggt atcaaaatga gcggcgtaag 1980

tacggtgtct cctgggcttt cactgatttt tccgacagta accagcgtca tgaagtctca 2040

cttgagggtc aggaacgcat ctggtcttca ccatatttga ttgtcgattt cctacccagt 2100

ctgtattacg aacaaaatac agaacacgat accccatact acaaccctat aaaaacgttc 2160

gatattgttc cggcatttga ggcaagccat ttgttatggc gaagctatga aaatagctgg 2220

gagcaaatat tcagcgcagg tgttggtgcc tcctggcaaa aacattatgg cacggatgtc 2280

gtcacccaac tcggctacgg gcaacgcatt agttggaatg acgtgattga tgctggcgca 2340

acgctacgct gggaaaaacg accttatgac ggtgacagag aacacaactt atacgttgaa 2400

ttcgatatga cattcagatt ttaaggataa atatgttacg taatggaaat aaatatctcc 2460

<210> 2
<211> 807
<212> PRT
<213> Escherichia coli

<400> 2

Met Tyr Ser Ser Ser Arg Lys Arg Cys Pro Lys Thr Lys Trp Ala Leu
1 5 10 15

Lys Leu Leu Thr Ala Ala Phe Leu Ala Ala Ser Pro Ala Ala Lys Ser

20 25 30

Ala Val Asn Asn Ala Tyr Asp Ala Leu Ile Ile Glu Ala Arg Lys Gly
35 40 45

Asn Thr Gln Pro Ala Leu Ser Trp Phe Ala Leu Lys Ser Ala Leu Ser
50 55 60

Asn Asn Gln Ile Ala Asp Trp Leu Gln Ile Ala Leu Trp Ala Gly Gln
65 70 75 80

Asp Lys Gln Val Ile Thr Val Tyr Asn Arg Tyr Arg His Gln Gln Leu
85 90 95

Pro Ala Arg Gly Tyr Ala Ala Val Ala Val Ala Tyr Arg Asn Leu Gln
100 105 110

Gln Trp Gln Asn Ser Leu Thr Leu Trp Gln Lys Ala Leu Ser Leu Glu
115 120 125

Pro Gln Asn Lys Asp Tyr Gln Arg Gly Gln Ile Leu Thr Leu Ala Asp
130 135 140

Ala Gly His Tyr Asp Thr Ala Leu Val Lys Leu Lys Gln Leu Asn Ser
145 150 155 160

Gly Ala Pro Asp Lys Ala Asn Leu Leu Ala Glu Ala Tyr Ile Tyr Lys
165 170 175

Leu Ala Gly Arg His Gln Asp Glu Leu Arg Ala Met Thr Glu Ser Leu
180 185 190

Pro Glu Asn Ala Ser Thr Gln Gln Tyr Pro Thr Glu Tyr Val Gln Ala
195 200 205

Leu Arg Asn Asn Gln Leu Ala Ala Ala Ile Asp Asp Ala Asn Leu Thr
210 215 220

Pro Asp Ile Arg Ala Asp Ile His Ala Glu Leu Val Arg Leu Ser Phe
225 230 235 240

Met Pro Thr Arg Ser Glu Ser Glu Arg Tyr Ala Ile Ala Asp Arg Ala
245 250 255

Leu Ala Gln Tyr Ala Ala Leu Glu Ile Leu Trp His Asp Asn Pro Asp

260 265 270

Arg Thr Ala Gln Tyr Gln Arg Ile Gln Val Asp His Leu Gly Ala Leu
275 280 285

Leu Thr Arg Asp Arg Tyr Lys Asp Val Ile Ser His Tyr Gln Arg Leu
290 295 300

Lys Lys Thr Gly Gln Ile Ile Pro Pro Trp Gly Gln Tyr Trp Val Ala
305 310 315 320

Ser Ala Tyr Leu Lys Asp His Gln Pro Lys Lys Ala Gln Ser Ile Met
325 330 335

Thr Glu Leu Phe Tyr His Lys Glu Thr Ile Ala Pro Asp Leu Ser Asp
340 345 350

Glu Glu Leu Ala Asp Leu Phe Tyr Ser His Leu Glu Ser Glu Asn Tyr
355 360 365

Pro Gly Ala Leu Thr Val Thr Gln His Thr Ile Asn Thr Ser Pro Pro
370 375 380

Phe Leu Arg Leu Met Gly Thr Pro Thr Ser Ile Pro Asn Asp Thr Trp
385 390 395 400

Leu Gln Gly His Ser Phe Leu Ser Thr Val Ala Lys Tyr Ser Asn Asp
405 410 415

Leu Pro Gln Ala Glu Met Thr Ala Arg Glu Leu Ala Tyr Asn Ala Pro
420 425 430

Gly Asn Gln Gly Leu Arg Ile Asp Tyr Ala Ser Val Leu Gln Ala Arg
435 440 445

Gly Trp Pro Arg Ala Ala Glu Asn Glu Leu Lys Lys Ala Glu Val Ile
450 455 460

Glu Pro Arg Asn Ile Asn Leu Glu Val Glu Gln Ala Trp Thr Ala Leu
465 470 475 480

Thr Leu Gln Glu Trp Gln Gln Ala Ala Val Leu Thr His Asp Val Val
485 490 495

Glu Arg Glu Pro Gln Asp Pro Gly Val Val Arg Leu Lys Arg Ala Val

500 505 510

Asp Val His Asn Leu Ala Glu Leu Arg Ile Ala Gly Ser Thr Gly Ile
515 520 525

Asp Ala Glu Gly Pro Asp Ser Gly Lys His Asp Val Asp Leu Thr Thr
530 535 540

Ile Val Tyr Ser Pro Pro Leu Lys Asp Asn Trp Arg Gly Phe Ala Gly
545 550 555 560

Phe Gly Tyr Ala Asp Gly Gln Phe Ser Glu Gly Lys Gly Ile Val Arg
565 570 575

Asp Trp Leu Ala Gly Val Glu Trp Arg Ser Arg Asn Ile Trp Leu Glu
580 585 590

Ala Glu Tyr Ala Glu Arg Val Phe Asn His Glu His Lys Pro Gly Ala
595 600 605

Arg Leu Ser Gly Trp Tyr Asp Phe Asn Asp Asn Trp Arg Ile Gly Ser
610 615 620

Gln Leu Glu Arg Leu Ser His Arg Val Pro Leu Arg Ala Met Lys Asn
625 630 635 640

Gly Val Thr Gly Asn Ser Ala Gln Ala Tyr Val Arg Trp Tyr Gln Asn
645 650 655

Glu Arg Arg Lys Tyr Gly Val Ser Trp Ala Phe Thr Asp Phe Ser Asp
660 665 670

Ser Asn Gln Arg His Glu Val Ser Leu Glu Gly Gln Glu Arg Ile Trp
675 680 685

Ser Ser Pro Tyr Leu Ile Val Asp Phe Leu Pro Ser Leu Tyr Tyr Glu
690 695 700

Gln Asn Thr Glu His Asp Thr Pro Tyr Tyr Asn Pro Ile Lys Thr Phe
705 710 715 720

Asp Ile Val Pro Ala Phe Glu Ala Ser His Leu Leu Trp Arg Ser Tyr
725 730 735

Glu Asn Ser Trp Glu Gln Ile Phe Ser Ala Gly Val Gly Ala Ser Trp

740 745 750

Gln Lys His Tyr Gly Thr Asp Val Val Thr Gln Leu Gly Tyr Gly Gln
755 760 765

Arg Ile Ser Trp Asn Asp Val Ile Asp Ala Gly Ala Thr Leu Arg Trp
770 775 780

Glu Lys Arg Pro Tyr Asp Gly Asp Arg Glu His Asn Leu Tyr Val Glu
785 790 795 800

Phe Asp Met Thr Phe Arg Phe
805

<210> 3
<211> 2031
<212> DNA
<213> Escherichia coli

<400> 3
ttaaggataa atatgttacg taatggaaat aaatatctcc tgatgctggt gagtataatt 60

atgctcaccg cgtgcattag ccagtcaaga acatcattta taccgccaca ggatcgcgaa 120

tctttactcg ccgagcaacc gtggccgcat aatggttttg tagcgatttc atggcataac 180

gttgaagacg aagctgccga ccagcgtttt atgtcagtgc ggacatcagc actgcgtgaa 240

caatttgcct ggctgcgcga gaacggttat caaccggtca gtattgctca aattcgtgaa 300

gcacatcgag gaggaaaacc gctaccggaa aaagctgtag tgctgacttt tgatgacggc 360

taccagagtt tttatacccg cgtcttccca attcttcagg ccttccagtg gcctgctgta 420

tgggcccccg tcggcagttg ggtcgatacg ccagcggata aacaagtaaa atttggcgat 480

gagttggtcg atcgagaata ttttgccacg tggcaacaag tgcgagaagt tgcgcgttcc 540

cggctcgttg agctcgcttc tcatacatgg aattctcact acggtattca ggctaatgcc 600

accggcagct tattgcctgt atatgtaaat cgtgcatatt ttactgacca cgcacggtat 660

gaaaccgcag cagaataccg ggaaagaatt cgtctggatg ctgtaaaaat gacggaatac 720

ctgcgtacaa aggttgaggt aaatccacac gtttttgttt ggccttatgg cgaagcgaat 780

ggcatagcga tagaggaatt aaaaaaactc ggttatgaca tgttcttcac ccttgaatca 840

ggtttggcaa atgcgtcgca attggattcc attccgcggg tattaatcgc caataatccc 900

tcattaaaag agtttgccca gcaaattatt accgtacagg aaaaatcacc acaacggata 960

atgcatatcg atcttgatta cgtttatgac gaaaacctcc agcaaatgga tcgcaatatt 1020

gatgtgctaa ttcagcgggt gaaagatatg caaatatcaa ccgtgtattt gcaggcattt 1080

gctgatcccg atggtgatgg gctggtcaaa gaggtctggt ttccaaatcg tttgctacca 1140

atgaaagcag atatttttag tcgggttgcc tggcaattac gtacccgctc aggtgtaaac 1200

atctatgcgt ggatgccggt attaagctgg gatttagatc ccacattaac gcgagtaaaa 1260

tacttaccaa caggggagaa aaaagcacaa attcatcctg aacaatatca ccgtctctct 1320

cctttcgatg acagagtcag agcacaagtt ggcatgttat atgaagatct tgccggacat 1380

gctgcttttg atggcatatt gttccacgat gatgctttgc tttcagatta tgaagatgcc 1440

agtgcaccgg ctatcacggc ttatcagcaa gcaggcttta gcgggagtct gagcgaaatt 1500

cgacaaaacc cggagcaatt taaacagtgg gcccgcttta aaagtcgtgc gttaactgac 1560

ttcactttag aacttagtgc gcgcgtaaaa gccattcgcg gtccacatat taaaactgca 1620

cgaaatattt ttgcacttcc ggtaatacaa cctgaaagtg aagcctggtt tgcacagaat 1680

tatgctgatt tcctaaaaag ctatgactgg accgctatta tggctatgcc ttatctggaa 1740

ggtgtcgcag aaaaatcggc tgaccaatgg ttaatacaat tgaccaatca aattaaaaac 1800

atccctcagg ctaaagacaa atctatttta gaattacagg cacaaaactg gcagaaaaat 1860

ggtcagcatc aggctatttc ttcgcaacaa ctcgctcact ggatgagcct attacaactg 1920

aatggagtga aaaactatgg ttattatccc gacaattttc tgcataacca acctgaaata 1980

gaccttattc gtcctgagtt ttcaacagcc tggtatccga aaaatgatta a 2031

<210> 4
<211> 672
<212> PRT
<213> Escherichia coli

<400> 4

Met Leu Arg Asn Gly Asn Lys Tyr Leu Leu Met Leu Val Ser Ile Ile
1 5 10 15

Met Leu Thr Ala Cys Ile Ser Gln Ser Arg Thr Ser Phe Ile Pro Pro
20 25 30

Gln Asp Arg Glu Ser Leu Leu Ala Glu Gln Pro Trp Pro His Asn Gly
35 40 45

Phe Val Ala Ile Ser Trp His Asn Val Glu Asp Glu Ala Ala Asp Gln
50 55 60

Arg Phe Met Ser Val Arg Thr Ser Ala Leu Arg Glu Gln Phe Ala Trp
65 70 75 80

Leu Arg Glu Asn Gly Tyr Gln Pro Val Ser Ile Ala Gln Ile Arg Glu
85 90 95

Ala His Arg Gly Gly Lys Pro Leu Pro Glu Lys Ala Val Val Leu Thr
100 105 110

Phe Asp Asp Gly Tyr Gln Ser Phe Tyr Thr Arg Val Phe Pro Ile Leu
115 120 125

Gln Ala Phe Gln Trp Pro Ala Val Trp Ala Pro Val Gly Ser Trp Val
130 135 140

Asp Thr Pro Ala Asp Lys Gln Val Lys Phe Gly Asp Glu Leu Val Asp
145 150 155 160

Arg Glu Tyr Phe Ala Thr Trp Gln Gln Val Arg Glu Val Ala Arg Ser
165 170 175

Arg Leu Val Glu Leu Ala Ser His Thr Trp Asn Ser His Tyr Gly Ile
180 185 190

Gln Ala Asn Ala Thr Gly Ser Leu Leu Pro Val Tyr Val Asn Arg Ala
195 200 205

Tyr Phe Thr Asp His Ala Arg Tyr Glu Thr Ala Ala Glu Tyr Arg Glu
210 215 220

Arg Ile Arg Leu Asp Ala Val Lys Met Thr Glu Tyr Leu Arg Thr Lys
225 230 235 240

Val Glu Val Asn Pro His Val Phe Val Trp Pro Tyr Gly Glu Ala Asn
245 250 255

Gly Ile Ala Ile Glu Glu Leu Lys Lys Leu Gly Tyr Asp Met Phe Phe
260 265 270

Thr Leu Glu Ser Gly Leu Ala Asn Ala Ser Gln Leu Asp Ser Ile Pro
275 280 285

Arg Val Leu Ile Ala Asn Asn Pro Ser Leu Lys Glu Phe Ala Gln Gln
290 295 300

Ile Ile Thr Val Gln Glu Lys Ser Pro Gln Arg Ile Met His Ile Asp
305 310 315 320

Leu Asp Tyr Val Tyr Asp Glu Asn Leu Gln Gln Met Asp Arg Asn Ile
325 330 335

Asp Val Leu Ile Gln Arg Val Lys Asp Met Gln Ile Ser Thr Val Tyr
340 345 350

Leu Gln Ala Phe Ala Asp Pro Asp Gly Asp Gly Leu Val Lys Glu Val
355 360 365

Trp Phe Pro Asn Arg Leu Leu Pro Met Lys Ala Asp Ile Phe Ser Arg
370 375 380

Val Ala Trp Gln Leu Arg Thr Arg Ser Gly Val Asn Ile Tyr Ala Trp
385 390 395 400

Met Pro Val Leu Ser Trp Asp Leu Asp Pro Thr Leu Thr Arg Val Lys
405 410 415

Tyr Leu Pro Thr Gly Glu Lys Lys Ala Gln Ile His Pro Glu Gln Tyr
420 425 430

His Arg Leu Ser Pro Phe Asp Asp Arg Val Arg Ala Gln Val Gly Met
435 440 445

Leu Tyr Glu Asp Leu Ala Gly His Ala Ala Phe Asp Gly Ile Leu Phe
450 455 460

His Asp Asp Ala Leu Leu Ser Asp Tyr Glu Asp Ala Ser Ala Pro Ala
465 470 475 480

Ile Thr Ala Tyr Gln Gln Ala Gly Phe Ser Gly Ser Leu Ser Glu Ile
485 490 495

Arg Gln Asn Pro Glu Gln Phe Lys Gln Trp Ala Arg Phe Lys Ser Arg
500 505 510

Ala Leu Thr Asp Phe Thr Leu Glu Leu Ser Ala Arg Val Lys Ala Ile
515 520 525

Arg Gly Pro His Ile Lys Thr Ala Arg Asn Ile Phe Ala Leu Pro Val
530 535 540

Ile Gln Pro Glu Ser Glu Ala Trp Phe Ala Gln Asn Tyr Ala Asp Phe
545 550 555 560

Leu Lys Ser Tyr Asp Trp Thr Ala Ile Met Ala Met Pro Tyr Leu Glu
565 570 575

Gly Val Ala Glu Lys Ser Ala Asp Gln Trp Leu Ile Gln Leu Thr Asn
580 585 590

Gln Ile Lys Asn Ile Pro Gln Ala Lys Asp Lys Ser Ile Leu Glu Leu
595 600 605

Gln Ala Gln Asn Trp Gln Lys Asn Gly Gln His Gln Ala Ile Ser Ser
610 615 620

Gln Gln Leu Ala His Trp Met Ser Leu Leu Gln Leu Asn Gly Val Lys
625 630 635 640

Asn Tyr Gly Tyr Tyr Pro Asp Asn Phe Leu His Asn Gln Pro Glu Ile
645 650 655

Asp Leu Ile Arg Pro Glu Phe Ser Thr Ala Trp Tyr Pro Lys Asn Asp
660 665 670

<210> 5
<211> 1380
<212> DNA
<213> Escherichia coli

<400> 5
aaaatgatta atcgcatcgt atcgtttttt atattatgtc tggtgttatg catacccctа 60

tgcgtagcgt actttcactc tggtgaactg atgatgaggt tcgttttctt ctggccgttt 120

tttatgtcca ttatgtggat tgttggcggc gtctatttct gggtctatcg tgaacgccac 180

tggccgtggg gagaaaacgc accagctccc cagttgaaag ataatccgtc tatctccatt 240

atcattccct gttttaatga ggagaaaaac gttgaggaaa ccatacacgc cgctttagca 300

cagcgttatg agaacattga agttattgcc gtaaatgacg gttcaacaga taaaacccgt 360

gccatcctgg atcgcatggc tgcacaaatt ccccatttgc gggtcattca tctggcgcaa 420

aaccagggga aagccattgc gcttaaaacc ggagctgccg cggcgaaaag tgaatatctg 480

gtgtgcattg atggcgatgc gttattagac cgcgatgcgg cggcatatat tgtggaaccg 540

atgttgtaca acccgcgtgt gggtgccgta accggtaatc ctcgtattcg aacacgttct 600

accctggtgg gtaaaattca ggttggcgag tattcctcaa ttattggttt gatcaagcga 660

acccagcgta tctatggaaa cgtatttacc gtttccggtg ttattgccgc atttcgtcgc 720

agcgccctgg cagaagtggg ttactggagt gacgatatga tcaccgaaga tattgatatt 780

agctggaagc tgcagttgaa tcagtggacg atttttttacg agccacgggc actgtgctgg 840

atattaatgc ctgaaacgtt aaaagggctg tggaaacagc gcctgcgctg ggctcagggc 900

ggtgcagaag tattcctcaa aaatatgaca aggttgtggc gcaaagaaaa ctttcgaatg 960

tggccgctgt tttttgaata ctgcctgacg acaatatggg ccttcacctg cctggtcggt 1020

ttcattattt acgcagtcca acttgccggt gtaccgttaa atattgaatt gacacatatc 1080

gctgcgacac atactgccgg aatattattg tgtacgttat gtttactgca atttattgtc 1140

agcctgatga tcgagaatcg ctatgagcat aatctgactt catcgctttt ctggattatt 1200

tggttcccgg ttatttttctg gatgctgagc ctggcaacga cattggtatc atttacacga 1260

gtcatgttga tgcctaaaaa gcaacgcgcc cgttgggtaa gtcccgatcg cgggattctg 1320

agaggttaat atgaacaatt taattattac gacccgacaa tcaccagtac gtttactggt 1380

<210> 6
<211> 445
<212> PRT
<213> Escherichia coli

<400> 6

Met Ile Asn Arg Ile Val Ser Phe Phe Ile Leu Cys Leu Val Leu Cys
1 5 10 15

Ile Pro Leu Cys Val Ala Tyr Phe His Ser Gly Glu Leu Met Met Arg
20 25 30

Phe Val Phe Phe Trp Pro Phe Phe Met Ser Ile Met Trp Ile Val Gly
35 40 45

Gly Val Tyr Phe Trp Val Tyr Arg Glu Arg His Trp Pro Trp Gly Glu
50 55 60

Asn Ala Pro Ala Pro Gln Leu Lys Asp Asn Pro Ser Ile Ser Ile Ile
65 70 75 80

Ile Pro Cys Phe Asn Glu Glu Lys Asn Val Glu Glu Thr Ile His Ala
85 90 95

Ala Leu Ala Gln Arg Tyr Glu Asn Ile Glu Val Ile Ala Val Asn Asp
100 105 110

Gly Ser Thr Asp Lys Thr Arg Ala Ile Leu Asp Arg Met Ala Ala Gln
115 120 125

```
Ile Pro His Leu Arg Val Ile His Leu Ala Gln Asn Gln Gly Lys Ala
    130                 135                 140

Ile Ala Leu Lys Thr Gly Ala Ala Ala Ala Lys Ser Glu Tyr Leu Val
145                 150                 155                 160

Cys Ile Asp Gly Asp Ala Leu Leu Asp Arg Asp Ala Ala Ala Tyr Ile
                165                 170                 175

Val Glu Pro Met Leu Tyr Asn Pro Arg Val Gly Ala Val Thr Gly Asn
            180                 185                 190

Pro Arg Ile Arg Thr Arg Ser Thr Leu Val Gly Lys Ile Gln Val Gly
        195                 200                 205

Glu Tyr Ser Ser Ile Ile Gly Leu Ile Lys Arg Thr Gln Arg Ile Tyr
    210                 215                 220

Gly Asn Val Phe Thr Val Ser Gly Val Ile Ala Ala Phe Arg Arg Ser
225                 230                 235                 240

Ala Leu Ala Glu Val Gly Tyr Trp Ser Asp Asp Met Ile Thr Glu Asp
                245                 250                 255

Ile Asp Ile Ser Trp Lys Leu Gln Leu Asn Gln Trp Thr Ile Phe Tyr
            260                 265                 270

Glu Pro Arg Ala Leu Cys Trp Ile Leu Met Pro Glu Thr Leu Lys Gly
        275                 280                 285

Leu Trp Lys Gln Arg Leu Arg Trp Ala Gln Gly Gly Ala Glu Val Phe
    290                 295                 300

Leu Lys Asn Met Thr Arg Leu Trp Arg Lys Glu Asn Phe Arg Met Trp
305                 310                 315                 320

Pro Leu Phe Phe Glu Tyr Cys Leu Thr Thr Ile Trp Ala Phe Thr Cys
                325                 330                 335

Leu Val Gly Phe Ile Ile Tyr Ala Val Gln Leu Ala Gly Val Pro Leu
            340                 345                 350

Asn Ile Glu Leu Thr His Ile Ala Ala Thr His Thr Ala Gly Ile Leu
        355                 360                 365
```

Leu Cys Thr Leu Cys Leu Leu Gln Phe Ile Val Ser Leu Met Ile Glu
370 375 380

Asn Arg Tyr Glu His Asn Leu Thr Ser Ser Leu Phe Trp Ile Ile Trp
385 390 395 400

Phe Pro Val Ile Phe Trp Met Leu Ser Leu Ala Thr Thr Leu Val Ser
405 410 415

Phe Thr Arg Val Met Leu Met Pro Lys Lys Gln Arg Ala Arg Trp Val
420 425 430

Ser Pro Asp Arg Gly Ile Leu Arg Gly Met Asn Asn Leu
435 440 445

<210> 7
<211> 30
<212> DNA
<213> Escherichia coli

<400> 7
tacagttaag tgtgttatcg gtgcagagcc 30

<210> 8
<211> 31
<212> DNA
<213> Escherichia coli

<400> 8
ctcaacgcct ggctgattaa accaactatt c 31

<210> 9
<211> 7500
<212> DNA
<213> Escherichia coli

<400> 9
atgtattcaa gtagcagaaa aaggtgcccg aaaaccaaat gggctttgaa acttcttact 60

gccgcatttt tagcagcgag tcccgcggcg aagagtgctg ttaataacgc ctatgatgca 120

ttgattattg aagctcgcaa gggtaatact cagccagctt tgtcatggtt tgcactaaaa 180

tcagcactca gcaataacca aattgctgac tggttacaga ttgccttatg ggccgggcaa 240

gataaacagg ttattaccgt ttacaaccgc taccgtcatc agcaattacc agcgcgtggt 300

tatgcagctg tcgccgtcgc ttatcgtaac ctgcaacaat ggcaaaactc gcttacactg 360

tggcaaaagg cgctctctct ggagccgcaa aataaggatt atcaacgggg acaaatttta 420

accctggcag atgctggtca ctatgatact gcgctggtta aacttaagca gcttaactct 480

ggagcaccgg acaaagccaa tttactcgca gaagcctata tctataaact ggcggggcgt 540

catcaggatg aattacgggc gatgacagag tcattacctg aaaatgcatc tacgcaacaa 600

tatcccacag aatacgtgca ggcattacgt aataatcaac ttgctgccgc gattgacgat 660

gccaatttaa cgccagatat tcgcgctgat attcatgccg aactggtcag actgtcgttt 720

atgcctacgc gcagtgaaag tgaacgttat gccattgccg atcgcgccct cgcccaatac 780

gctgcattag aaattctgtg gcacgataac ccagaccgca ctgcccagta ccagcgtatt 840

caggttgatc atcttggcgc gttattaact cgcgatcgtt ataaagacgt tatttctcac 900

tatcagcgat taaaaaagac ggggcaaatt attccgccct gggggcaata ttgggttgca 960

tcggcttatc tcaaagatca tcagccgaaa aaagcacagt caataatgac cgagctcttt 1020

tatcacaagg agaccattgc cccggattta tccgatgaag aacttgcgga tctcttttac 1080

agccacctgg agagtgaaaa ttatccgggc gcgctaactg tcacccaaca taccattaat 1140

acttcgccgc ctttccttcg gttaatgggc acgcctacga gcatcccgaa tgatacctgg 1200

ttacaggggc attcgtttct ctcaaccgta gcaaaatata gtaatgatct tcctcaggct 1260

gaaatgacag ccagagagct tgcttataac gcaccaggaa atcagggact gcgcattgat 1320

tacgcgagtg tgttacaagc ccgcggttgg cctcgtgcag cagaaaatga attaaaaaaa 1380

gcagaagtga tcgagccacg taatattaat ctggaggttg aacaagcctg gacagcatta 1440

acgttacaag aatggcagca ggcagctgtc ttaacgcacg atgttgtcga acgtgaaccg 1500

caagatcccg gcgttgtacg attaaaacgt gcggttgatg tacataatct tgcagagctt 1560

cgtatcgctg gctcaacagg aattgatgcc gaaggcccgg atagtggtaa acatgatgtc 1620

gacttaacca ccatcgttta ttcaccaccg ctgaaggata actggcgcgg ttttgctgga 1680

ttcggttatg ccgatggaca atttagcgaa ggaaaaggga ttgttcgcga ctggcttgcg 1740

ggtgttgagt ggcggtcacg taatatctgg ctcgaggcag agtacgctga acgcgttttc 1800

aatcatgagc ataaacccgg cgcgcgcctg tctggctggt atgattttaa tgataactgg 1860

cgtattggtt cgcaactgga acgcctctct caccgcgttc cattacgggc aatgaaaaat 1920

ggtgttacag gcaacagtgc tcaggcttat gttcgctggt atcaaaatga gcggcgtaag 1980

tacggtgtct cctgggcttt cactgatttt tccgacagta accagcgtca tgaagtctca 2040

cttgagggtc aggaacgcat ctggtcttca ccatatttga ttgtcgattt cctacccagt 2100

ctgtattacg aacaaaatac agaacacgat accccatact acaaccctat aaaaacgttc 2160

gatattgttc cggcatttga ggcaagccat ttgttatggc gaagctatga aaatagctgg 2220

gagcaaatat tcagcgcagg tgttggtgcc tcctggcaaa aacattatgg cacggatgtc 2280

gtcacccaac tcggctacgg gcaacgcatt agttggaatg acgtgattga tgctggcgca 2340

acgctacgct gggaaaaacg accttatgac ggtgacagag aacacaactt atacgttgaa 2400

ttcgatatga cattcagatt ttaaggataa atatgttacg taatggaaat aaatatctcc 2460

tgatgctggt gagtataatt atgctcaccg cgtgcattag ccagtcaaga acatcattta 2520

taccgccaca ggatcgcgaa tctttactcg ccgagcaacc gtggccgcat aatggttttg 2580

tagcgatttc atggcataac gttgaagacg aagctgccga ccagcgtttt atgtcagtgc 2640

ggacatcagc actgcgtgaa caatttgcct ggctgcgcga gaacggttat caaccggtca 2700

gtattgctca aattcgtgaa gcacatcgag gaggaaaacc gctaccggaa aaagctgtag 2760

tgctgacttt tgatgacggc taccagagtt tttatacccg cgtcttccca attcttcagg 2820

ccttccagtg gcctgctgta tgggcccccg tcggcagttg ggtcgatacg ccagcggata 2880

aacaagtaaa atttggcgat gagttggtcg atcgagaata ttttgccacg tggcaacaag 2940

tgcgagaagt tgcgcgttcc cggctcgttg agctcgcttc tcatacatgg aattctcact 3000

acggtattca ggctaatgcc accggcagct tattgcctgt atatgtaaat cgtgcatatt 3060

ttactgacca cgcacggtat gaaaccgcag cagaataccg ggaaagaatt cgtctggatg 3120

ctgtaaaaat gacggaatac ctgcgtacaa aggttgaggt aaatccacac gtttttgttt 3180

ggccttatgg cgaagcgaat ggcatagcga tagaggaatt aaaaaaactc ggttatgaca 3240

tgttcttcac ccttgaatca ggtttggcaa atgcgtcgca attggattcc attccgcggg 3300

tattaatcgc caataatccc tcattaaaag agtttgccca gcaaattatt accgtacagg 3360

aaaaatcacc acaacggata atgcatatcg atcttgatta cgtttatgac gaaaacctcc 3420

agcaaatgga tcgcaatatt gatgtgctaa ttcagcgggt gaaagatatg caaatatcaa 3480

ccgtgtattt gcaggcattt gctgatcccg atggtgatgg gctggtcaaa gaggtctggt 3540

ttccaaatcg tttgctacca atgaaagcag atatttttag tcgggttgcc tggcaattac 3600

gtacccgctc aggtgtaaac atctatgcgt ggatgccggt attaagctgg gatttagatc 3660

ccacattaac gcgagtaaaa tacttaccaa caggggagaa aaaagcacaa attcatcctg 3720

aacaatatca ccgtctctct cctttcgatg acagagtcag agcacaagtt ggcatgttat 3780

atgaagatct tgccggacat gctgcttttg atggcatatt gttccacgat gatgctttgc 3840

tttcagatta tgaagatgcc agtgcaccgg ctatcacggc ttatcagcaa gcaggcttta 3900

gcgggagtct gagcgaaatt cgacaaaacc cggagcaatt taaacagtgg gcccgcttta 3960

aaagtcgtgc gttaactgac ttcactttag aacttagtgc gcgcgtaaaa gccattcgcg 4020

gtccacatat taaaactgca cgaaatattt ttgcacttcc ggtaatacaa cctgaaagtg 4080

aagcctggtt tgcacagaat tatgctgatt tcctaaaaag ctatgactgg accgctatta 4140

tggctatgcc ttatctggaa ggtgtcgcag aaaaatcggc tgaccaatgg ttaatacaat 4200

tgaccaatca aattaaaaac atccctcagg ctaaagacaa atctatttta gaattacagg 4260

cacaaaactg gcagaaaaat ggtcagcatc aggctatttc ttcgcaacaa ctcgctcact 4320

ggatgagcct attacaactg aatggagtga aaaactatgg ttattatccc gacaattttc 4380

tgcataacca acctgaaata gaccttattc gtcctgagtt ttcaacagcc tggtatccga 4440

aaaatgatta atcgcatcgt atcgtttttt atattatgtc tggtgttatg catacccta 4500

tgcgtagcgt actttcactc tggtgaactg atgatgaggt tcgttttctt ctggccgttt 4560

tttatgtcca ttatgtggat tgttggcggc gtctatttct gggtctatcg tgaacgccac 4620

tggccgtggg gagaaaacgc accagctccc cagttgaaag ataatccgtc tatctccatt 4680

atcattccct gttttaatga ggagaaaaac gttgaggaaa ccatacacgc cgctttagca 4740

cagcgttatg agaacattga agttattgcc gtaaatgacg gttcaacaga taaaacccgt 4800

gccatcctgg atcgcatggc tgcacaaatt ccccatttgc gggtcattca tctggcgcaa 4860

aaccagggga aagccattgc gcttaaaacc ggagctgccg cggcgaaaag tgaatatctg 4920

gtgtgcattg atggcgatgc gttattagac cgcgatgcgg cggcatatat tgtggaaccg 4980

atgttgtaca acccgcgtgt gggtgccgta accggtaatc ctcgtattcg aacacgttct 5040

accctggtgg gtaaaattca ggttggcgag tattcctcaa ttattggttt gatcaagcga 5100

acccagcgta tctatggaaa cgtatttacc gtttccggtg ttattgccgc atttcgtcgc 5160

agcgccctgg cagaagtggg ttactggagt gacgatatga tcaccgaaga tattgatatt 5220

agctggaagc tgcagttgaa tcagtggacg atttttttacg agccacgggc actgtgctgg 5280

atattaatgc ctgaaacgtt aaaagggctg tggaaacagc gcctgcgctg ggctcagggc 5340

ggtgcagaag tattcctcaa aaatatgaca aggttgtggc gcaaagaaaa ctttcgaatg 5400

tggccgctgt tttttgaata ctgcctgacg acaatatggg ccttcacctg cctggtcggt 5460

ttcattattt acgcagtcca acttgccggt gtaccgttaa atattgaatt gacacatatc 5520

gctgcgacac atactgccgg aatattattg tgtacgttat gtttactgca atttattgtc 5580

agcctgatga tcgagaatcg ctatgagcat aatctgactt catcgctttt ctggattatt 5640

tggttcccgg ttatttttctg gatgctgagc ctggcaacga cattggtatc atttacacga 5700

gtcatgttga tgcctaaaaa gcaacgcgcc cgttgggtaa gtcccgatcg cgggattctg 5760

agaggttaat atgaacaatt taattattac gacccgacaa tcaccagtac gtttactggt 5820

tgattatgtt gccacaacca tcttgtggac attatttgcg ttgttcatat tcttattcgc 5880

catggatctg ctgacgggtt attactggca aagcgaggcc agaagccgac ttcagttcta 5940

ttttttgctg gcagtggcga atgccgtcgt gttaattgtc tgggcgctgt acaataagct 6000

gcgttttcaa aaacagcagc atcatgcagc ctaccaatat acgccgcaag aatatgcaga 6060

gagcttagca atacctgatg agctctatca gcaactacaa aaaagccaca ggatgagcgt 6120

acacttcacc agccaggggc aaataaaaat ggttgtttca gaaaaagcgc tagtccgggc 6180

ataaacaccc aaaacaaagc ccggttcgcc cgggctctgc accgataaca cacttaactg 6240

taggcatgca gcgtacgttg gcaaagtgcc gaacgtacgc agtcctcttt accgaaccgg 6300

acgatcccaa ccatttcatc ttcttcgaaa cgttccagcg cgtcacttaa tccggagcac 6360

acgccgcgag gcaaatcgca ttgcgtgata tcaccgttga cgataaccgt cacgttctcc 6420

ccgaggcggg ttaaaaacat tttcatttgc gcggcagtca cattctgcgc ctcgtcaaga 6480

atgacgactg cattttcaaa ggtacgtcca cgcatatagg cgaacggcgc aatttccacc 6540

ttccctattt ccggtcgcag gcagtactgc ataaaggaag cccctaagcg ccggaccagc 6600

acgtcgtaga ccgggcgaaa atagggagca aacttttctg cgatatctcc aggtaagaag 6660

ccaagatctt catcggcttg cagaactgga cgggtgacga taatcctgtc gacatcctta 6720

tgtatcaggg cctctgccgc ttttgctgcg ctgatccagg tttttccgca cccggcttcg 6780

cccgtggcga atatcagctg cttactctca atagccttca gatagtgcaa ttgcgcttca 6840

tttcgcgcga ggatgggcga agtatcgcga ctgtcgcggg ccataccaat ggcttctacg 6900

ccgcccatct gcacaagcga ggtgaccgat tcttcttcac gctgcttatg gctgcgcgaa 6960

tcccgtctca gcacacgttt tgcctcgcga cgagctttga tcactgcttt ttgtcttccc 7020

atggagagca ccttgagttg tttgtattca tcacacgcgc cgttggcagc gcgattatgc 7080

gcacgaacat cagagggttg gcttccttgt aagccatagt ttgcttttgg ataaaatgcc 7140

gaaaaacggc tacgcgcacc gtttacggcg tcggtaacac atgaaaagaa aggatgaggt 7200

tgaaaatgca aagtgacgag atgactaccg gaggagaaaa ctccgcgagt ggtggcgcgt 7260

tgattatcta aaacatgtcc agtacaggac gttaccatcc gcgatctcca tagtgactga 7320

ctatcactgc cgggaacttc cgctgctact taataagtac aacagatctc gcatttattg 7380

caacaatata tttacttata tttaactata aaacaccatt tcagtgacat tagtttctac 7440

tggaaagatg acagagtgat gacagtgatg aaaaaagctg tgtgctttca gcaggatttg 7500